Optimal inversion of hard X-ray bremsstrahlung spectra

I. SVD analysis

M. Piana[1] and J.C. Brown[2]

[1] INFM and Dipartimento di Fisica, Università di Genova, Genova, Italy
[2] Department of Physics and Astronomy, University of Glasgow, Glasgow, UK

Abstract. The instability properties of bremsstrahlung spectrum inversion to yield source electron spectra in flares are examined in terms of singular value decomposition (SVD) for the cases of thin- and thick-target models for different bremsstrahlung cross-sections. It is shown that, for the Kramers and Bethe-Heitler cross-sections, analytic expressions can be obtained for evaluation of the Gram matrix. These allow quantitative comparison of the instability of inversion for each case and hence assessment of the accuracy and resolution of the electron source function recovery. Based on this analysis we find that: (a) the solution instability is worse for the Bethe-Heitler than for the Kramers approximation; care must therefore be taken in real data analysis to use the most accurate possible cross-section to avoid over- or under-estimating the maximum precision and resolution achievable in the recovered electron spectrum; (b) instability is worse in recovering thick-target injection spectra than thin-target mean source spectra, as expected because of the double deconvolution involved in the former; (c) such SVD analysis should form an important part of future analysis of high resolution flare data such as from the HESSI Mission.

Key words: X-rays: general — methods: analytical; data analysis

1. Introduction and formulation of the integral equations

We consider the analysis of solar flare hard X-rays (HXR) in the energy range from around ten to a few hundred keV emitted by collisional bremsstrahlung of high energy electrons on ions of the ambient plasma. We assume that the electron distribution function is isotropic and the hydrogen dominated plasma is fully ionized (cf. discussion of the latter by Brown et al. 1998). It is well-known that HXRs carry detailed information about flare electrons and, in particular, the shape of the spectra are related to electron acceleration and energy transport. Careful consideration of techniques for optimal recovery of such information is particularly timely with NASA's planned launch of the HESSI flare mission, the first to carry a high resolution HXR spectrometer. Through the 1960s and 70s only low spectral HXR resolution ($\approx 20 - 30\%$) was available and without any spatial resolution. The 1980s and 90s saw the first HXR imaging (SMM, Hinotori and YohKoh missions) and the separate advent of Ge detector applications yielding HXR resolution to $\approx 1\%$. The latter allowed the first possibility of genuine HXR spectral inversion while the former gave the first indications of the HXR source structure. HESSI is the first instrument to combine both these functions and will enable new insight into the spectral distibution of source electrons as functions of position and time, facts of major interest to flare modelling. To obtain the maximum benefit from these data will require the development of optimal methods for deconvolving the observed photon spectra as well as the spatial information obtained by the Fourier imaging technique involved.

In the case of an optically thin source of collisional bremsstrahlung, the total rate of photon emission $g(\epsilon)$ per unit photon energy ϵ is given by

$$g(\epsilon) = \int_V n(\boldsymbol{r}) \int_\epsilon^\infty F(E, \boldsymbol{r}) Q(\epsilon, E) \mathrm{d}E \mathrm{d}\boldsymbol{r} \qquad (1)$$

where V is the volume of the emitting region, $n(\boldsymbol{r})$ is the density of the ions in the plasma, $F(E, \boldsymbol{r})$ is the local electron flux spectrum, E is the electron energy and $Q(\epsilon, E)$ is the bremsstrahlung cross-section. We note that g represents the output at the sun while actual measurements are of the output at the earth but since only the shape of the spectrum and the relative errors are considered here, multiplicative constants are irrelevant.

In a non-thermal atmosphere, Eq. (1) can be simplified according to two different models. If the electron lifetimes inside the source region are long compared to the observation time, then $F(E, \boldsymbol{r})$ is not modified by energy

loss and reflects the instantaneous flux spectrum of high-energy electrons. In this non-thermal thin-target model (Brown 1971) by defining the average ion density

$$\overline{n} = \frac{1}{V} \int_V n(\boldsymbol{r})\mathrm{d}\boldsymbol{r} \qquad (2)$$

and the average electron distribution function

$$f(E) = \frac{1}{\overline{n}V} \int_V n(\boldsymbol{r})F(E,\boldsymbol{r})\mathrm{d}\boldsymbol{r} \qquad (3)$$

equation (1) can be simplified to

$$g(\epsilon) = \int_\epsilon^\infty f(E)Q(\epsilon,E)\mathrm{d}E. \qquad (4)$$

If the lifetimes of the electrons in the source region are short compared to the observation time, then electrons have to be continuously injected at a rate $F_0(E_0)$ s^{-1} per unit injection energy E_0. Within this non-thermal thick-target framework the photon flux is given by (Brown 1971)

$$g(\epsilon) = \int_\epsilon^\infty F_0(E_0) \int_\epsilon^{E_0} \frac{Q(\epsilon,E)}{R(E)}\mathrm{d}E\mathrm{d}E_0 \qquad (5)$$

where $R(E)$ represents the average collisional energy loss rate of an electron of energy E per unit plasma column density. Equation (5) can be written in the form of a Volterra integral equation of the first kind

$$g(\epsilon) = \int_\epsilon^\infty \frac{Q(\epsilon,E)}{R(E)}G(E)\mathrm{d}E \qquad (6)$$

by defining

$$G(E) = \int_E^\infty F_0(E_0)\mathrm{d}E_0. \qquad (7)$$

Both these models involve linear integral equations of the first kind, where the data are represented by the continuous bremsstrahlung spectrum and the solutions are energetic electron distribution functions - the electron average distribution function $f(E)$ or the injection rate $F_0(E_0)$. To find these functions from Eqs. (4)-(7) an explicit form for the bremsstrahlung cross-section $Q(\epsilon,E)$ must be chosen. For typical applications ($\epsilon \simeq 10-100$ keV) non-relativistic expressions are often adequate. The simplest form is given by Kramers' formula (Koch & Motz 1959)

$$Q(\epsilon,E) = \frac{Q_0}{\epsilon E} \qquad (8)$$

where for hydrogen

$$Q_0 = \frac{8}{3}\frac{r_0^2}{137}mc^2 \qquad (9)$$

with r_0 the classical radius and m the rest mass of the electron respectively. A more accurate approximation is given by the Bethe-Heitler formula

$$Q(\epsilon,E) = \frac{Q_0}{\epsilon E} \log \frac{1+\sqrt{1-\frac{\epsilon}{E}}}{1-\sqrt{1-\frac{\epsilon}{E}}}. \qquad (10)$$

However relativistic and semi-relativistic corrections have also been introduced (Gould 1980; Haug 1997) and, in general, the sensitivity of the solutions of Eqs. (4) or (5) to variations in the cross-section is an interesting and important problem (see results below).

Inversion of Eq. (4), with the Bethe-Heitler cross-section only, has been addressed in (Piana 1994). In that paper the ill-posed nature of the problem of determining $f(E)$ from knowledge of $g(\epsilon)$ was discussed and Tikhonov regularized versions of the solution were computed by using the singular system of the integral operator in the case of discrete data. Applications to noisy simulated and real spectra were also considered. In the present paper we want to generalize the singular value decomposition (SVD) approach to the case of non-thermal bremsstrahlung for both thick and thin targets and for both forms of Q above. Our aim is to provide a quantitative basis, in terms of SVD analysis, for evaluation of the differences resulting from the two cross-section approximations, and between the thick- and thin-target cases, in respect of the ill-posedness of the problems, i.e., how accurately the source functions can be recovered from noisy data in each case. In Sect. 4 we compare these results with the analogous ones for the more ill-posed problem of thermal bremsstrahlung inversion. In a forthcoming paper, devoted to applications, we will use the SVD analysis to invert a real HXR spectrum emitted during a solar flare.

2. SVD analysis for the thin-target model

Let us consider Eq. (4) with a general form for the cross-section $Q(\epsilon,E)$. In practical applications the photon spectrum is known only at discrete values of the photon energy and the equation we have to deal with is

$$g(\epsilon_n) = \int_{\epsilon_n}^\infty Q(\epsilon_n,E)f(E)\mathrm{d}E \qquad n=1,\ldots,N. \qquad (11)$$

Actually detectors provide integrals of the spectrum over finite ranges of photon energy and a more realistic version of the discrete equation is represented by Eq. (48) in (Piana 1994). However in that paper it was shown that as far as the inversion is concerned, the difference between the use of point data and binned data is not significant, though we need to allow for the effect of binning on photon count statistics.

We assume that the function f belongs to the Hilbert space $L^2(0,\infty)$ of the functions φ such that $\int_0^\infty |\varphi(x)|^2\mathrm{d}x < \infty$. The choice of this source space is for mathematical reasons, though it is not physically unrealistic to assume that the electron distribution functions are square integrable. We introduce the integral operator $A: L^2(0,\infty) \rightarrow Y$ defined by

$$(Af)_n = \int_{\epsilon_n}^\infty Q(\epsilon_n,E)f(E)\mathrm{d}E \qquad n=1,\ldots,N. \qquad (12)$$

Y is a finite dimension vector space equipped with the inner product

$$(\boldsymbol{g},\boldsymbol{h})_Y = \sum_{n,m=1}^N g_n w_{nm} h_m. \qquad (13)$$

The most suitable choice for the weights w_{nm} depends on the sampling adopted (Bertero et al. 1984). For example, in the case of the uniform sampling,

$$\epsilon_n = \epsilon_1 + d(n-1) \qquad n = 1, \ldots, N \tag{14}$$

with d constant, the typical choice is

$$w_n = d \qquad n = 1, \ldots, N \tag{15}$$

while, in the case of the geometric sampling,

$$\epsilon_n = \epsilon_1 \Delta^{n-1} \qquad n = 1, \ldots, N \tag{16}$$

with Δ a constant, it is usually

$$w_n = (\log \Delta)\epsilon_n \qquad n = 1, \ldots, N. \tag{17}$$

Therefore the linear inverse problem with discrete data (11) we are interested in can be described by the equation

$$\boldsymbol{g} = Af. \tag{18}$$

We note that Riesz's lemma (Reed & Simon 1972) allows us to describe the action of the integral operator by use of the scalar product $(\cdot, \cdot)$ in $L^2(0, \infty)$, i.e.

$$(Af)_n = (f, \phi_n) \qquad n = 1, \ldots, N \tag{19}$$

with

$$\phi_n(E) = \begin{cases} 0 & E \leq \epsilon_n \\ Q(\epsilon_n, E) & E > \epsilon_n \end{cases} \tag{20}$$

for $n = 1, \ldots, N$. We define the adjoint operator $A^* : Y \to L^2(0, \infty)$ of A by

$$(A^*\boldsymbol{g}, f) = (\boldsymbol{g}, Af)_Y. \tag{21}$$

Then the singular system of the operator A is the set of triples $\{\sigma_n; u_n, \boldsymbol{v}_n\}_{n=1}^N$ which solves the set of problems (Bertero et al. 1985)

$$Au_n = \sigma_n \boldsymbol{v}_n; \qquad A^*\boldsymbol{v}_n = \sigma_n u_n \qquad n = 1, \ldots, N. \tag{22}$$

The real non-negative numbers σ_n, $n = 1, \ldots, N$ form a non-increasing sequence and are called singular values, the functions u_n, $n = 1, \ldots, N$ are called singular functions and the vectors $\boldsymbol{v}_n$, $n = 1, \ldots, N$ are called singular vectors.

Knowledge of the singular system of the operator A is extremely useful for the solution of the linear inverse problem with discrete data (18). In particular:

- the singular system gives information about the degree of instability of the solution of the inverse problem;
- it represents a basic tool for the computation of regularization algorithms providing approximate and stable solutions of the problem;
- from the distribution of the zeros of the singular functions it is possible to infer information about the resolution obtainable from the data.

Now we want to describe these three properties of the singular system in detail.

If $Q(\epsilon, E)$ is the Kramers or Bethe-Heitler cross-section, the linear inverse problem of determining the electron distribution function f from the photon spectrum g when f and g are related by Eq. (4) is ill-posed in the sense of Hadamard (Hadamard 1923). A problem is said well-posed in the sense of Hadamard if:

- the solution of the problem exists for each data function.
- the solution is unique;
- the solution depends continuously on the data.

If at least one of these three requirements is not satisfied, the problem is ill-posed in the sense of Hadamard. In the case of the inverse problem with discrete data (18) the solution is clearly not unique - i.e., infinitely many functions f yield precisely the same data vector $\boldsymbol{g}$, though it can be shown (Bertero 1989) that the problem of determining the solution among these which has minimum norm (the so called generalized solution) is well-posed. However, though the generalized solution of (18) is unique for precisely given $\boldsymbol{g}$, it is unstable to noise on $\boldsymbol{g}$ - that is the problem is "ill-conditioned". It can be shown that if $\delta\boldsymbol{g}$ and δf are respectively the error on the data and the error on the (generalized) solution, the inequality

$$\frac{\|\delta f\|}{\|f\|} \leq C(A)\frac{\|\delta\boldsymbol{g}\|_Y}{\|\boldsymbol{g}\|_Y} \tag{23}$$

holds. Here $\|\cdot\|$ and $\|\cdot\|_Y$ denote respectively the norm in $L^2(0, \infty)$, i.e.

$$\|\varphi\| = \sqrt{\int_0^\infty |\varphi(x)|^2 \mathrm{d}x} \tag{24}$$

and the norm in Y, i.e.

$$\|\boldsymbol{g}\|_Y = \sqrt{\sum_{n,m=1}^N g_n w_{nm} g_m}. \tag{25}$$

The number $C(A)$, which depends on the explicit form of the operator A, is called the condition number and clearly represents a reliable indicator of the (potential) instability of the problem, for if $C(A)$ is large from Eq. (23) a small relative error on the data may imply a significant relative error on the solution. It can be shown that the relation

$$C(A) = \frac{\sigma_1}{\sigma_N} \tag{26}$$

holds, which shows that the singular value spectrum of the operator carries explicit information about the instability of the problem.

It is well-known that the numerical instability of the solution of problem (18) can be reduced by applying regularization. In the case of linear regularization methods a family of linear operators $R_r : Y \to L^2(0, \infty)$, with r the so-called regularization parameter, is introduced and an approximate stable solution of the problem is looked for in the set of regularized solutions $f_r = R_r\boldsymbol{g}$. The best of these regularized solutions is obtained by choosing an optimum value of the regularization parameter r and to this aim "ad hoc" criteria have been formulated (Davies 1992). Examples of linear regularization methods (Bertero 1989) are represented by Tikhonov's formulation, where

$$R_\lambda = (A^*A + \lambda I)^{-1}A^* \tag{27}$$

and $r = \lambda$ is real positive and by Landweber's iterative formulation, where

$$R_k = \tau \sum_{j=0}^{k-1} (I - \tau A^* A)^j A^* \tag{28}$$

with

$$0 < \tau < \frac{2}{\|A\|_{\sup}^2} \tag{29}$$

and $r = k$ is a positive integer (here $\|\cdot\|_{\sup}$ denotes the canonical operator norm). The computation of the singular value decomposition of the operator A is basic for fast application of both these methods since the action of the operators R_λ and R_k can be described in terms of the singular system. In fact it can be shown that

$$f_\lambda(E) = R_\lambda \boldsymbol{g} = \sum_{n=1}^N \frac{\sigma_n}{\sigma_n^2 + \lambda} (\boldsymbol{g}, \boldsymbol{v}_n)_Y u_n(E) \tag{30}$$

and

$$f_k(E) = R_k \boldsymbol{g} = \tau \sum_{n=1}^N \frac{[1 - (1 - \tau \sigma_n^2)^k]}{\sigma_n} (\boldsymbol{g}, \boldsymbol{v}_n)_Y u_n(E). \tag{31}$$

We note that also the computation of the typical criteria for the choice of the regularization parameter is simplified if the singular system of A is known (Bertero 1989).

Knowledge of the behaviour of the singular functions also gives information about the resolution achievable for given data and a given source space (in the present case we have assumed $L^2(0, \infty)$ as the source space). In linear inverse problem theory the concept of resolution limit is deeply related to a concept called *optimum choice of the experiment*, meaning that measurements are made in such a way as to give the maximum possible source function resolution consistent with the limits imposed by noise. As already pointed out, the minimum norm solution, or generalized solution, of problem (18) exists and is unique; its expression in terms of the singular system is

$$f(E) = \sum_{n=1}^N \frac{1}{\sigma_n} (\boldsymbol{g}, \boldsymbol{v}_n)_Y u_n(E). \tag{32}$$

If we increase the number of experimental points N, the discretized problem (18) becomes closer and closer to the continuous problem (4) and, as a consequence of the ill-posedness of this problem and of the presence of noise on the data, the generalized solution becomes more and more unstable. In other words, by adding more points, we do not add more information and, even in the limiting case of an infinite-dimensional problem, from a given data function, only a finite amount of information can be derived. Obviously this finite amount of information can be obtained with some finite number, say N_0, of points. Terminating sum (32) at $N = N_0$ is optimal in that it recovers the full information available while stabilizing the solution against random noise variations. We could make the solution more stable by reducing N further, but this has the price of reducing the information on resolution

in the solution. If the N_0-th singular function has zeros, from Eq. (32) the generalized solution cannot contain details in intervals smaller than the distance ΔE between adjacent zeros and the ratio $\Delta E/E$ may be adopted as a measure of the resolution corresponding to N_0. In general the optimum number N_0 of points is not known and has to be chosen according to how we decide to trade-off resolution versus instability. Let us then consider the case of a noisy vector $\boldsymbol{g}$ of N data points, with $N > N_0$, and let us apply, for example, Tikhonov's algorithm (30), in which high order terms in (32) are increasingly suppressed rather than eliminated. In the case of expansion (30) the number of singular functions which significantly contribute to the restoration of the source function is in general less than N. We denote this number again by N_0. A possible criterion for choosing N_0 is that the relative error due to the truncation of the expansion (30) must be just smaller than the measured relative error on the photon spectrum. Then the N_0-th singular function is the last one giving a significant contribution to the regularized solution defined in (30). If ΔE is the distance between the zeros of the N_0-th singular function so chosen, $\Delta E/E$ is the corresponding resolution limit (with "higher resolution" meaning smaller $\Delta E/E$).

The computation of the singular system of the operator A is based on the observation (cf. Eq. (22)) that the singular values and the singular vectors are given respectively by the square root of the eigenvalues and by the eigenvectors of the matrix AA^*. Furthermore it can be proven that (Bertero et al. 1985)

$$AA^* = G^T W \tag{33}$$

where W is the weight matrix whose mn entry is defined by

$$W_{mn} = w_m \delta_{mn} \tag{34}$$

and G is the Gram matrix whose mn entry is defined by

$$G_{mn} = (\phi_m, \phi_n) \tag{35}$$

with the functions ϕ_n, $n = 1, \ldots, N$ defined as in (20).

If we adopt the Kramers cross-section the explicit form of the functions ϕ_n is given by

$$\phi_n(E) = \begin{cases} 0 & E \leq \epsilon_n \\ \frac{1}{\epsilon_n E} & E > \epsilon_n \end{cases} \tag{36}$$

and the mn entry of the Gram matrix is, for $m \geq n$

$$G_{mn} = \frac{1}{\epsilon_m^2 \epsilon_n} \tag{37}$$

and for $m < n$

$$G_{mn} = G_{nm}. \tag{38}$$

If we adopt the Bethe-Heitler approximation, the explicit form of the functions ϕ_n is given by (Piana 1994)

$$\phi_n(E) = \begin{cases} 0 & E \leq \epsilon_n \\ \frac{1}{\epsilon_n E} \log \frac{1 + \sqrt{1 - \frac{\epsilon_n}{E}}}{1 - \sqrt{1 - \frac{\epsilon_n}{E}}} & E > \epsilon_n \end{cases} \tag{39}$$

Table 1. Condition numbers for the thin-target model in the case of the Kramers (K) and the Bethe-Heitler (B − H) approximations for different numbers N of sampling points. The photon energy range is $\epsilon_1 = 10$ keV, $\epsilon_N = 100$ keV and the sampling is geometric

	$N = 10$	$N = 25$	$N = 50$	$N = 100$
K	40	114	243	509
B − H	183	922	2915	8846

and the mn entry of the Gram matrix is, for $m > n$:

$$G_{mn} = \frac{4}{(\epsilon_n \epsilon_m)^{\frac{3}{2}}} \log \frac{1 + \sqrt{\frac{\epsilon_n}{\epsilon_m}}}{1 - \sqrt{\frac{\epsilon_n}{\epsilon_m}}} -$$

$$- \frac{2}{\epsilon_n (\epsilon_m)^2} \log \frac{\epsilon_n}{\epsilon_m} +$$

$$+ \frac{2}{\epsilon_n (\epsilon_m)^2} \left(1 + \frac{\epsilon_m}{\epsilon_n} \right) \log \left(1 - \frac{\epsilon_n}{\epsilon_m} \right); \tag{40}$$

for $m = n$:

$$G_{mm} = \frac{1}{(\epsilon_m)^3} 8\log 2; \tag{41}$$

for $m < n$:

$$G_{mn} = G_{nm}. \tag{42}$$

In Table 1 we present the values of the condition number for $N = 10, 25, 50, 100$ geometrically sampled points in the energy range $\epsilon_1 = 10$ keV, $\epsilon_N = 100$ keV in the case of the Kramers and Bethe-Heitler cross-sections. From these results the following facts can be deduced:

- for both cross-sections the condition number increases with the number of sampling points. This behaviour can be explained by observing that in the limit $N \to \infty$ the finite rank operator defined in (19) tends to the integral operator defined by the right hand side of Eq. (4). This operator is compact and its singular spectrum tends to zero. More roughly speaking, if the number of sampling points increases, the discrete ill-conditioned problem becomes more and more similar to the continuous problem which is ill-posed and whose ill-conditioning is formally infinite;
- the condition number is systematically bigger if the Bethe-Heitler cross-section is adopted instead of the Kramers one because the former is a smoother kernel. This means that in the case of the Bethe-Heitler cross-section the inverse problem of determining the electron distribution function from the photon spectrum is more unstable in the presence of data noise. The factor of 10 increase in $C(A)$ for $N = 25$ data points, for example, implies that the data accuracy needed for the recovery of the electron spectrum with prescribed accuracy is 10 times higher when the Bethe-Heitler formula is used.

In Fig. 1 the first four singular functions in the cases of the Kramers and Bethe-Heitler cross-sections' are plotted

for $N = 25$ geometrically sampled points and energy range between $\epsilon_1 = 10$ keV and $\epsilon_N = 100$ keV. These singular functions are computed by using the formula

$$u_n(E) = \frac{1}{\sigma_n} \sum_{j=1}^{N} w_j (\boldsymbol{v}_n)_j \phi_j(E). \tag{43}$$

We note that the singular functions in the Kramers case are discontinuous at $E = \epsilon_n$ $n = 1, \ldots, N$ (they are linear combinations of the functions ϕ_n in (36) which are discontinuous at these energies). Furthermore, for both approximations, the singular function of order n is characterized by n zeros, approximately displaced according to a geometric progression of ratio p. For a given data vector $\boldsymbol{g}$ (and for a given source space) the value of p in the last significant singular function (i.e., giving a significant contribution to Eq. (30)), describes the achievable resolution limit. Therefore the value of this limit explicitly depends on the level of noise affecting the data. However it is possible to roughly foresee the difference in the resolution limit achievable for the Kramers or the Bethe-Heitler approximations. For example, let us take $N = 25$ points geometrically sampled between $\epsilon_1 = 10$ keV and $\epsilon_N = 100$ keV. In the Kramers approximation, the condition number for 25 points is $C(A) \simeq 114$ and the value of p for the 25-th singular function is $p \simeq 1.1$. It follows that the zeros of the 25-th singular function occur at

$$E_k = E_1 p^{k-1} \qquad k = 1, \ldots, 25 \tag{44}$$

so that

$$\frac{\Delta E_k}{E_k} = p - 1 \simeq 0.1 \tag{45}$$

is the resolution limit (10%). In the Bethe-Heitler approximation the condition number for 25 points is larger ($C(A) \simeq 922$). However, if we reduce the number of points used in the Bethe-Heitler case, we find that the ratio σ_1/σ_{11} is approximately equal to the condition number in the Kramers cross-section case for all 25 points (i.e., σ_1/σ_{25} in the Kramers cross-section case). For the eleventh Bethe-Heitler singular function the value of p is $p \simeq 1.2$ and in this case $\Delta E_k/E_k \simeq 0.2$. From this we conclude that the resolution in the Kramers cross-section approximation is double that for the Bethe-Heitler, for the same solution accuracy.

3. SVD analysis for the thick-target model

We consider now the case of the thick-target model. Within this framework the determination of the injection rate $F_0(E_0)$ from the photon spectrum can be accomplished in two ways. A first possibility is represented by explicitly solving the inner integral at the right hand side of Eq. (5). If $\tilde{Q}(\epsilon, E_0)$ is the result of this integration, F_0 can be determined by solving the linear inverse problem with discrete data

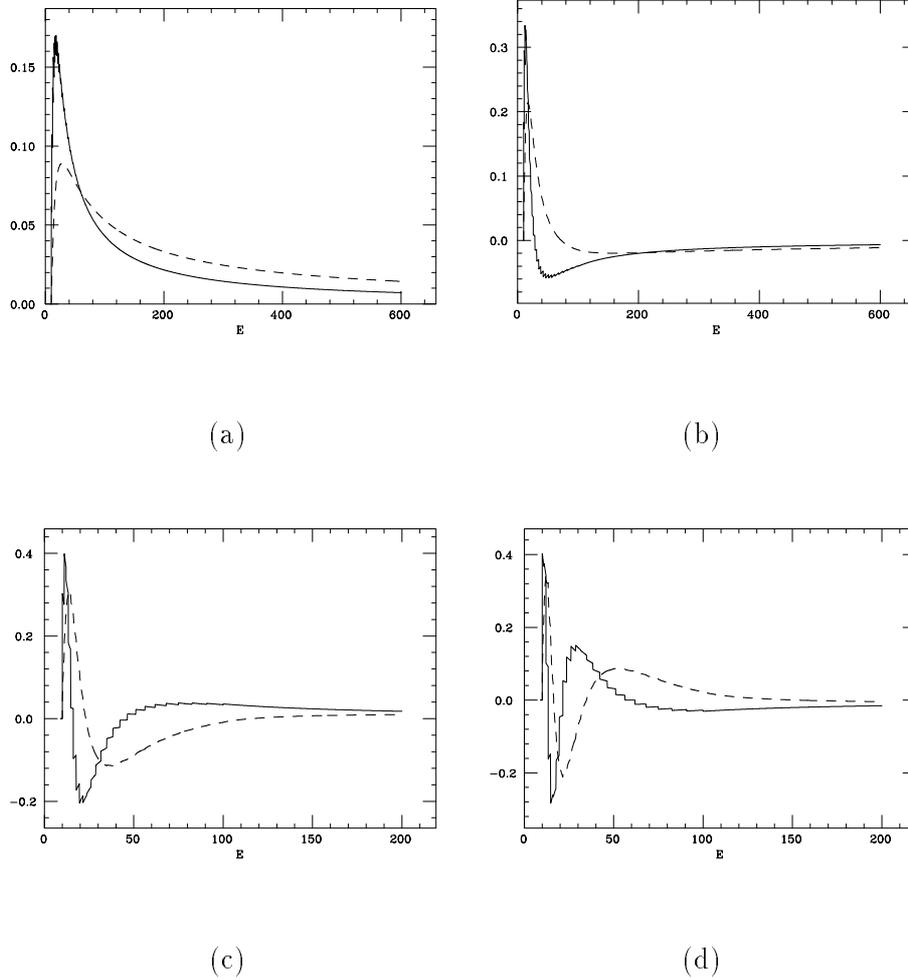$$\boldsymbol{g} = A_0 F_0 \tag{46}$$

(a)

(b)

(c)

(d)

Fig. 1. Singular functions for the thin-target model in the case of the Kramers (solid) and the Bethe-Heitler (dashes) cross-section:
a) first singular function; **b)** second singular function; **c)** third singular function; (4) fourth singular function

with A_0 defined by

$$(A_0 F_0)(\epsilon_n) = \int_{\epsilon_n}^{\infty} \tilde{Q}(\epsilon_n, E_0) F_0(E_0) \mathrm{d}E_0 \qquad (47)$$

with $n = 1, \ldots, N$. Otherwise F_0 can be obtained by consecutively solving the two linear ill-posed inverse problems represented by the discrete version of Eq. (6) and by Eq. (7).

We note that for both these approaches it is necessary to choose an explicit form for the average energy loss rate $R(E)$. In the case of a collisional cold but ionized thick target characterized only by Coulomb losses we have

$$R(E) = \frac{K}{E} \qquad (48)$$

with $K = 2\pi e^4 \Lambda$; e is the electron charge and Λ is a parameter which depends on the Debye length and which can be considered constant in the energy range typical of this application.

If the determination of F_0 is addressed by solving problem (5) the singular value decomposition can be performed

in the case of the Kramers and the Bethe-Heitler approximations by computing the Gram matrix. In the case of the Kramers cross-section we find

$$\tilde{Q}(\epsilon, E_0) = \frac{Q_0}{K} \frac{E_0 - \epsilon}{\epsilon} \qquad (49)$$

and in order to deal with square integrable integral kernels we consider the linear inverse problem with discrete data

$$g(\epsilon_n) = \frac{1}{\epsilon_n} \int_{\epsilon_n}^{\infty} G_\alpha(E_0) \frac{E_0 - \epsilon_n}{E_0^\alpha} \mathrm{d}E_0 \qquad (50)$$

with $G_\alpha(E_0) = F_0(E_0) E_0^\alpha$ (all the multiplicative constants have been put equal to one). If $\alpha > 3/2$ the functions

$$\phi_n(E_0) = \begin{cases} 0 & E \le \epsilon_n \\ \frac{1}{\epsilon_n} \frac{E_0 - \epsilon_n}{E_0^\alpha} & E > \epsilon_n \end{cases} \qquad (51)$$

are in $L^2(0, \infty)$ for all $n = 1, \ldots, N$. For $n \ge m$ the nm entry of the Gram matrix is given by

$$G_{nm} = \frac{1}{\epsilon_n \epsilon_m} \left[\frac{1}{2\alpha - 3} \frac{1}{\epsilon_n^{2\alpha - 3}} - \right.$$
$$\left. - (\epsilon_n + \epsilon_m) \frac{1}{2\alpha - 2} \frac{1}{\epsilon_n^{2\alpha - 2}} + \right.$$

Table 2. Condition numbers for the thick-target model in the case of the Kramers (K) and the Bethe-Heitler (B − H) approximations for different numbers N of sampling points. The photon energy range is $\epsilon_1 = 10$ keV, $\epsilon_N = 100$ keV and the sampling is geometric

	$N = 10$	$N = 25$	$N = 50$	$N = 100$
K	248	2143	9950	43709
B − H	868	11647	78071	494383

$$+ \; \epsilon_n \epsilon_m \frac{1}{2\alpha - 1} \frac{1}{\epsilon_n^{2\alpha-1}} \Bigg] \tag{52}$$

and for $n < m$

$$G_{nm} = G_{mn}. \tag{53}$$

If the Bethe-Heitler cross-section is assumed, we have

$$g(\epsilon_n) = \frac{1}{\epsilon_n} \int_{\epsilon_n}^{\infty} G_0(E_0) \Bigg[\frac{1}{E_0^{\alpha-1}} \log \frac{1 + \sqrt{1 - \frac{\epsilon_n}{E_0}}}{1 - \sqrt{1 - \frac{\epsilon_n}{E_0}}} +$$

$$- \frac{1}{E_0^{\alpha-1}} \sqrt{1 - \frac{\epsilon_n}{E_0}} -$$

$$- \frac{1}{2} \frac{\epsilon_n}{E_0^{\alpha}} \log \frac{1 + \sqrt{1 - \frac{\epsilon_n}{E_0}}}{1 - \sqrt{1 - \frac{\epsilon_n}{E_0}}} \Bigg] \, dE_0. \tag{54}$$

We assume again $\alpha > 3/2$ so that the functions

$$\phi_n(E_0) = \begin{cases} 0 & E_0 \leq \epsilon_n \\ \frac{1}{\epsilon_n} \Bigg[\frac{1}{E_0^{\alpha 1}} \log \frac{1 + \sqrt{1 - \frac{\epsilon_n}{E_0}}}{1 - \sqrt{1 - \frac{\epsilon_n}{E_0}}} - \\ - \frac{1}{E_0^{\alpha-1}} \sqrt{1 - \frac{\epsilon_n}{E_0}} - & E > \epsilon_n \\ - \frac{\epsilon_n}{2E_0} \log \frac{1 + \sqrt{1 - \frac{\epsilon_n}{E_0}}}{1 - \sqrt{1 - \frac{\epsilon_n}{E_0}}} \Bigg] \end{cases} \tag{55}$$

are square integrable; in this case the entries of the Gram matrix are computed by using numerical integration.

The condition numbers for the problems (50) and (54) are presented in Table 2 and in Fig. 2 we plot the first four singular functions in the case of $N = 25$ geometrically sampled points in the energy range $\epsilon_1 = 10$ keV, $\epsilon_N = 100$ keV. As in the case of the thin-target model, if the Bethe-Heitler cross-section is adopted the instability of the problem is greater and the resolution limit achievable is poorer for prescribed solution accuracy.

4. The thermal case

For completeness and comparison we also give here (based on Piana et al. 1995) the parallel analysis of the thermal inverse problem where the emission is assumed to come from an optically thin thermal source (Craig & Brown 1976). If $n_e(\boldsymbol{r})$, $T(\boldsymbol{r})$ are the electron density and temperature at the point $\boldsymbol{r}$ in the source region and $F(\epsilon, T(\boldsymbol{r}))$ is

the photon spectral distribution function for an isothermal source element, the photon spectrum is given by

$$g(\epsilon) = \int_V F(\epsilon, T(\boldsymbol{r})) n_e^2(\boldsymbol{r}) d\boldsymbol{r}. \tag{56}$$

From this equation, once the Kramers' cross-section is adopted and via suitable changes of variable (Craig & Brown 1976) we have

$$g(\epsilon) = \frac{1}{\epsilon} \int_0^{\infty} \frac{\xi(T)}{\sqrt{kT}} \exp\left(-\frac{\epsilon}{kT}\right) dT \tag{57}$$

where k is the Boltzmann constant and $\xi(T)$ is the differential emission measure defined by

$$\xi(T) = \int_{S_T} \frac{n_e^2}{|\nabla T|} dS_T \tag{58}$$

with S_T a constant temperature surface.

Equation (57) can be easily transformed into the inverse Laplace transform problem

$$g(\epsilon) = \frac{1}{\epsilon} \int_0^{\infty} \eta(y) \exp(-\epsilon y) dy \tag{59}$$

with $y = 1/kT$ and

$$\eta(y) = y^{-\frac{3}{2}} \xi\left(\frac{1}{ky}\right) \tag{60}$$

(all the multiplicative constants have been put equal to one). A linear inverse problem with discrete data

$$\boldsymbol{g} = L\eta \tag{61}$$

can also be written in this case, with L defined by

$$(L\eta)(\epsilon_n) = \frac{1}{\epsilon_n} \int_0^{\infty} \eta(y) \exp(-\epsilon_n y) dy \quad n = 1, \ldots, N. \tag{62}$$

If we introduce the $L^2(0, \infty)$ functions

$$\phi_n(y) = \frac{1}{\epsilon_n} \exp(-\epsilon_n y) \quad n = 1, \ldots, N \tag{63}$$

the Gram matrix in this case is given by

$$G_{mn} = \frac{1}{\epsilon_m \epsilon_n} \frac{1}{\epsilon_m + \epsilon_n} \tag{64}$$

for $n, m = 1, \ldots, N$.

In Table 3 we compare the condition numbers for $N = 10, 25, 50, 100$ geometrically sampled points (the energy range is again $\epsilon_1 = 10$ Kev, $\epsilon_N = 100$ keV) for the thermal case and the non-thermal models (thin- and thick-target) when the Kramers cross-section is used (the Kramers' cross-section is the one adopted to derive Eq. (57)). As one can see, the Laplace inversion problem is much more unstable than the non-thermal problems.

5. Conclusions

Our SVD and condition number analysis of the flare bremsstrahlung spectrum inversion problem has established the following important results:
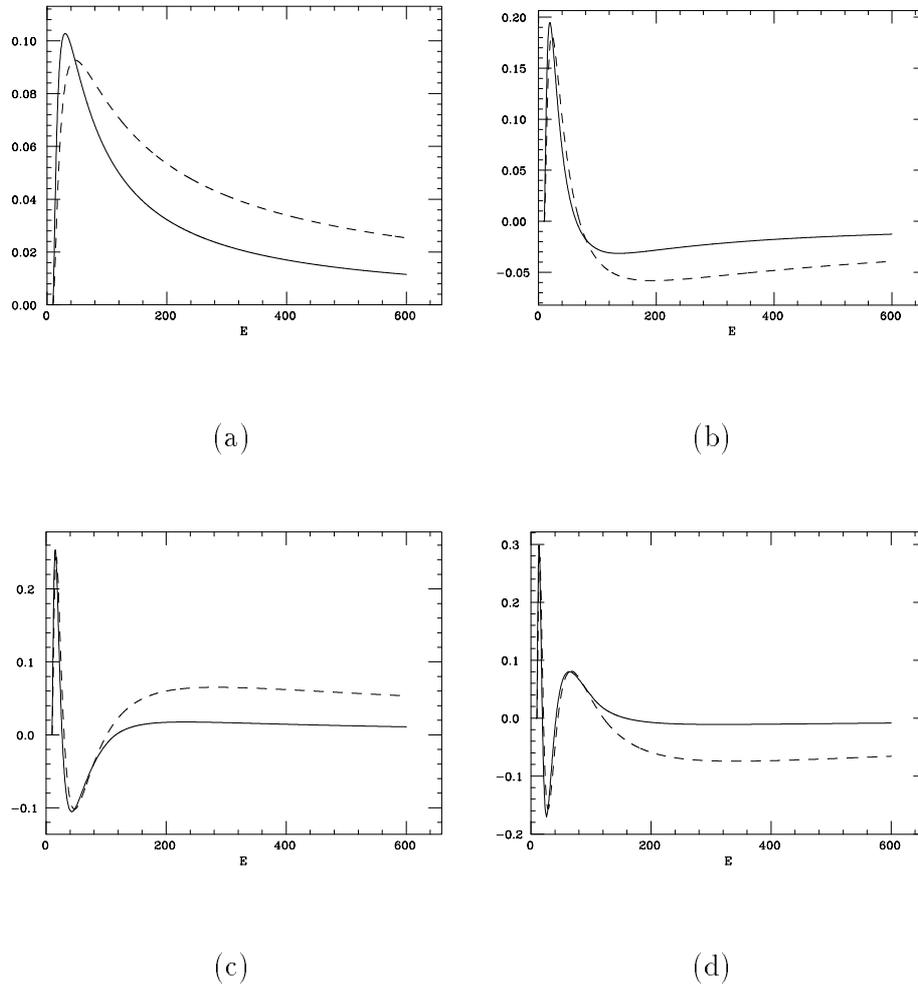
(a)

(b)

(c)

(d)

Fig. 2. Singular functions for the thick-target model in the case of the Kramers (solid) and the Bethe-Heitler (dashes) cross-section: **a)** first singular function; **b)** second singular function; **c)** third singular function; (4) fourth singular function

Table 3. Condition numbers for the thin-target, thick-target and thermal models when the Kramers cross-section is adopted, for different numbers N of sampling points. The photon energy range is $\epsilon_1 = 10$ keV, $\epsilon_N = 100$ keV and the sampling is geometric

	$N = 10$	$N = 25$	$N = 50$	$N = 100$
Thin	40	114	243	509
Thick	248	2143	9950	43709
Thermal	$5.7\,10^5$	$9.6\,10^8$	$2.5\,10^9$	$5.4\,10^9$

- analytic expressions can be obtained for computing the singular functions and values for thick- and thin-target problems with both the Kramers and the Bethe-Heitler cross-sections. These make it straightforward to assess quantitatively the accuracy and resolution achievable in the electron spectrum solution in each case;
- the condition number (error magnification) as usual increases with the number of recovery points. It is

considerably larger for the Bethe-Heitler than for the Kramers cross-section. Insofar as the Bethe-Heitler expression is physically more accurate, this means that if we used the Kramers cross section for numerical inversions we would overestimate the accuracy of the recovered solution - i.e. undersmooth it (alternatively we can say that in finding a regularized solution we will need to smooth the Kramers case less than the Bethe-Heitler one). In fact, however, there are aspects of the Kramers expression which are physically *more* accurate than the Bethe-Heitler. Therefore, given that our results show the ill-posedness to be sensitive to the cross-section, it will be worthwhile to pursue further our present analysis by numerical analysis of the inversion properties of more precise cross-sections;
- the ill-conditioning of the thick-target inversion is considerably worse than for the thin-target, regardless of cross-section, so that the accuracy with which we can recover the injected flux spectrum from prescribed data is a great deal lower than that of the mean

bremsstrahlung source (thin-target) spectrum. This is as expected because of the double convolution (cross-section and electron path) involved between the thick-target injection spectrum and the HXR data. It is also consistent with the second derivative analytic solution given by Brown & Emslie (1989) and the heuristic numerical solution by Johns & Lin (1992);

– in the case of the thermal model the inversion of the photon spectrum to determine the differential emission measure is much more ill-conditioned than the inverse problems considered in the two non-thermal models.

References

Bertero M., 1989, in: Hawkes P.W. (ed.) Advances in Electronics and Electron Physics 75. Academic Press

Bertero M., De Mol C., Pike E.R., 1984, Proc. R. Soc. Lond. A 393, 51

Bertero M., De Mol C., Pike E.R., 1985, Inverse Problems 1, 301

Brown J.C., 1971, Solar Phys. 18, 489

Brown J.C., Emslie A.G., 1989, Astrophys. J. 331, 554

Brown J.C., McArthur G.A., Barrett R.K., McIntosh S.W., Emslie A.G., 1998, Solar Phys., (in press)

Craig I.J.D., Brown J.C., 1976, A&A 49, 239

Craig I.J.D., Brown J.C., 1986, Inverse Problems in Astronomy, Hilger

Davies A.M., 1992, in Bertero M. and Pike E.R. (eds.) Inverse Problems in Scattering and Imaging, Hilger

Gould R.J., 1980, ApJ 238, 1026

Hadamard J., 1923, Lectures on Cauchy's Problems in Linear Partial Differential Equations. Yale University Press

Haug E., 1997, A&A 326, 417

Johns C.M., Lin R.P., 1992, Solar Phys. 137, 121

Koch H.W., Motz J.W., 1959, Rev. Mod. Phys. 31, 920

Piana M., 1994, A&A 288, 949

Piana M., Brown J.C., Thompson A.M., 1995, Solar Phys. 156, 315

Reed M., Simon B., 1972, Methods of Modern Mathematical Physics I. Academic Press